SEMARAK ILMU
PUBLISHING
202103268166(003316878-P)

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

A Framework for Detection of Face Mask Using Deep Learning Approach with Internet of Things (IoT)

Pravat Kumar Rautaray[1], Binod Kumar Pattanayak[1,*], Mihir Narayan Mohanty[2], Bibhuprasad Mohanty[2]

[1] Department of Computer Science & Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan, deemed to be University, Bhubaneswar, 751030 Odisha, India
[2] Department of Electronics and Communication Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan, deemed to be University, Bhubaneswar, 751030 Odisha, India

ARTICLE INFO

Article history:
Received 29 May 2023
Received in revised form 30 November 2023
Accepted 4 March 2024
Available online 25 June 2024

Keywords:
IoMT; Transfer learning; Convolutional neural networks

ABSTRACT

A wide range of diseases have become more common in modern human society. Without proper care, the mortality rate consequently rises, resulting in a loss of worth. Humans now place a high importance on genuinely caring for their health and wealth. A facemask covering nose and mouth is one of the most effective ways to protect against infection and to spread the corona virus. This safeguard rule is applied by almost all governments. For automation, we have developed an approach based on deep learning to detect the mask. The approach has been extended to an Internet of Things (IoT) based framework that can be an element of a smart city to keep people safe. Since health safety is a major challenge, this paper proposes an automatic detection process. A Convolutional neural network with transfer learning is used for the detection. This model is included in an IoT architecture for automation. The results of testing the approach are have been observed to be excellent in terms of accuracy. In addition, the module for IoT works well, as verified in the study, and also appears to be useful for the Internet of Medical Things.

1. Introduction

Internet of Things (IOT) plays a vital role in all fields from agriculture to healthcare. The Internet of Things (IoT) can be used to automate and monitor greenhouses. IoT sensors can be installed in greenhouses to collect data about the environment, including climate, light exposure, carbon dioxide, and soil properties proposed by Chin *et al.*, [22]. Novel corona virus (nCoV) spreads silently and cannot be identified early. It requires that one should be alert by self. This virus causes sickness that exhibits symptoms of common cold flue showing the kinds of infections such as "Middle East Respiratory Syndrome (MERS)" that results in "Severe Acute Respiratory Syndrome (SARS)" [1]. In the month of December 2019, the corona virus's (referred to as COVID-19) earliest infected patient

* *Corresponding author.*
E-mail address: binodpattanayak@soa.ac.in

https://doi.org/10.37934/araset.48.2.2844

was found in Huang of China. Since then, COVID-19 has spread worldwide and become a pandemic. Due to this pandemic, people all over the world are facing difficulties like job loss, finance related difficulties, and occasionally death. Since then, numerous deaths have occurred, and this number keeps rising. The most significant symptoms emerging from a patient infected with COVID-19 virus, as identified by the World Health Organization(WHO) are such as loss of taste, loss of smell, dry cough, diarrhoea, fever tiredness and people have been instructed to follow specific precautionary measures such as washing hands regularly, avoiding social contact and wearing a face mask in crowded places, ceasing from frequently touching mouth, nose as well as eyes to prevent oneself from being infected by this deadly virus. The spread of the said virus can be prevented by wearing a mask and maintaining strict social distance [4]. Unfortunately, these rules are not properly followed by the people, which is the main reason for the spread of the virus. In an effort to stop the spread of the Corona virus, these irresponsible individuals who are not adhering to the rules, should be noticed by the appropriate authorities.

Several approaches to object detection have been presented. Applications in medicine make extensive use of the approaches such as deep learning. Detection of face mask detection can benefit from such approaches. Additionally, a metropolitan area with numerous IOT sensors for data collection is considered a smart city.

In this paper the proposed system has been designed for an effective autonomous process of finding out whether a particular person is wearing a face mask or is without it, thereby reducing the physical efforts. This system basically relies on the fundamental principles of Image processing and Internet of Things (IOT). If any person is not using a mask, then the automated system has the ability to inform the competent authority in a smart city network. Using real-time footage from the city's CCTV cameras in a variety of public areas makes this possible. Facial pictures from the real time videos are removed and then, these images are utilized to recognize the face mask that is wrapped on the face of an individual. By virtue of using the "Convolutional Neural Network (CNN)" learning algorithm, the features of the image are extracted and then, can be used for parameters' learning of various hidden layers incorporated in it. Whenever the system will find no mask on the face of an individual, then such information would be transmitted via the Internet of Things module to corresponding competent authority to take necessary actions. Utilizing the proposed framework, it is possible to guarantee that the law will be properly implemented against individuals who are not adhering to essential health rules during this pandemic.

The rest of this paper has been structured as detailed below. Recent contributions from various authors in the available literature on detection of face mask are elaborated in Section 2. Section 3 covers the proposed here model for fostering the said framework that has been portrayed. Section 4 investigates the outcomes of the proposed in this paper model and Section 5 concludes the paper with focus on probable extension of this work.

2. Related Work

Proposed model used deep learning, Keras, OpenCV, and TensorFlow for detection of face mask. Here for classification, they have used a framework MobilenetV2 architecture. As it is light weight in nature, this model can be used in various embedded devices such as like Raspberry pi, NVIDIA Jetson Nano in order for conducting detection of face mask in real-time. This model gives an accuracy of 92.64 and F1 Score 93 percentage [3]. Here the author proposed a model where they used MobilenetV2 as classifier and after training the model is saved for future use. Then the same model was implemented in the video. The face detection algorithm works after the video reads frame by frame. In case of detection of a face, it continues to the next step. Here the model achieved 92 percent of accuracy [4]. An alarm system-based real-time facemask recognition model was presented here by the authors with the implementation of deep learning approaches using CNN. The image will be collected with the help of a camera installed and passed these images using a trained model that is further sent to a Raspberry pi device. Here, the authors have used VGG16 as the pre-trained model which could achieve an estimated accuracy equal to 96 percents. Militante *et al.*, [5] made comparisons between three classifiers; the first one is the Support Vector Machine (SVM), the second one is the Decision Tree (DT)and the third one is the proposed as well as developed by the authors themselves. And, as claimed by them, it was observed that the proposed model could achieve 91.11 percent of accuracy that is assumed to be enhanced further. Hussain *et al.,* [6] created one proposed model where they used YOLO v2 along with ResNet-50 for the purpose of classification and claimed to have achieved an average precession of 81 percent. The authors here have conducted a comparison and further fine-tuned taking into consideration various optimizers [7]. It should be noted that the details of captured images may be successfully restored with the help of Super-resolution (SR) networks. The concepts of transfer learning and Auto-encoder have recently been incorporated into SR networks for the purpose of performance enhancement. Before image segmentation or classification, image processing techniques also make use of SR networks in order to achieve better performance by Das, A *et al.,* [9] and also to reconstruct images in order for achieving higher resolution thereby restoring details. In addition to it, such networks extend a viable option for enhancing the performance of condition identification for facemask-wearing individuals thereby improving the accuracy of classification of images significantly, particularly while working with a dataset that comprises of the images of very bad quality. In order for improvement of the accuracy level during classification of facial images in particular, a SR network can be used along with a "SR classification network (SRCNet)". This proposed model produced 98.70 percent of accuracy. Bosheng *et al.,* [8] investigated on how to avoid spread of corona virus we can sanitize the infected area by using deep learning [10]. The diseases can be detected by virtue of Predicting the patients affected by COVID making use of "Gated Recurrent Unit Neural Networks" Das *et al.,* [11] and Patra [12]. Bhoi *et al*., [13] demonstrated that their research able to control and make people aware of healthcare by implementing IoT technology.

3. The Proposed Model

An automated face mask detector will be used in the proposed method to screen individuals who do not wear face masks. CCTV cameras will be used to keep an eye on this entire community area. The system will process through the pictures that the cameras take. If a person is identified without wearing a mask on his/her face, their photos will be sent to the authorities for immediate action. And a person found to have body temperature more than normal then immediately that message will be transmitted to nearby Ambulance. The newly developed model has been illustrated in the block diagram depicted in Figure 1.

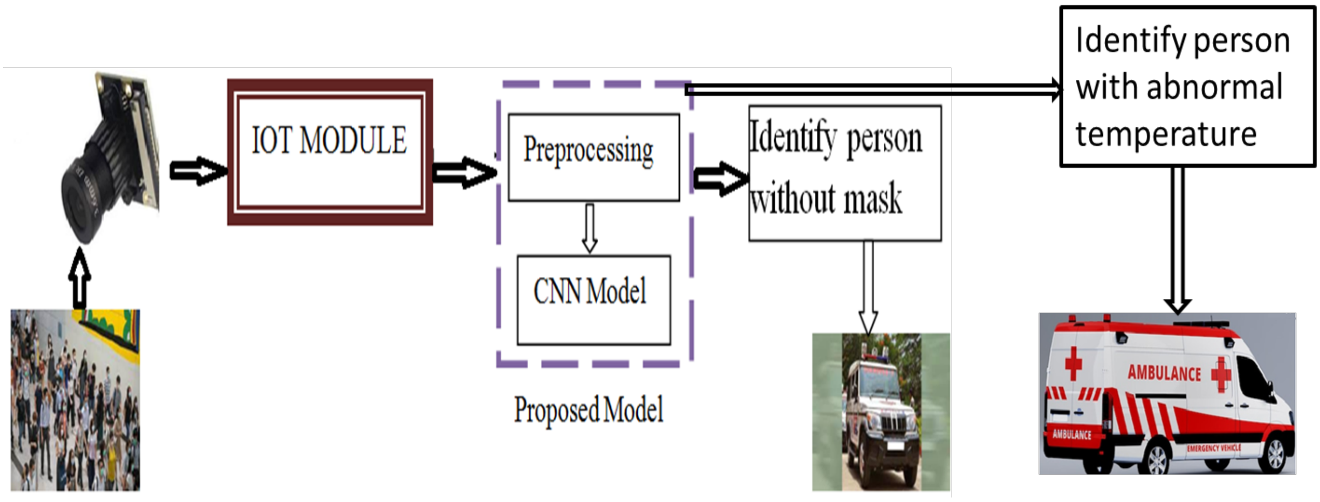

Fig. 1. Model Architecture

All blocks of the proposed here model is clarified as follows. Figure 2 shows the persons wearing a face mask as well as without a face mask.

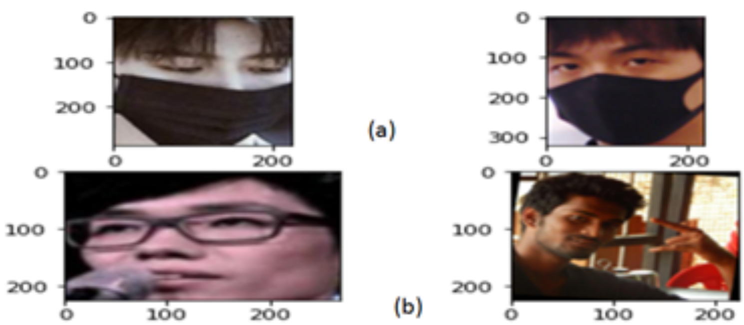

Fig. 2. (a) Two sample of person with mask. (b) Two samples of person without mask

Figure 3 defines the sequence of steps of the operation.

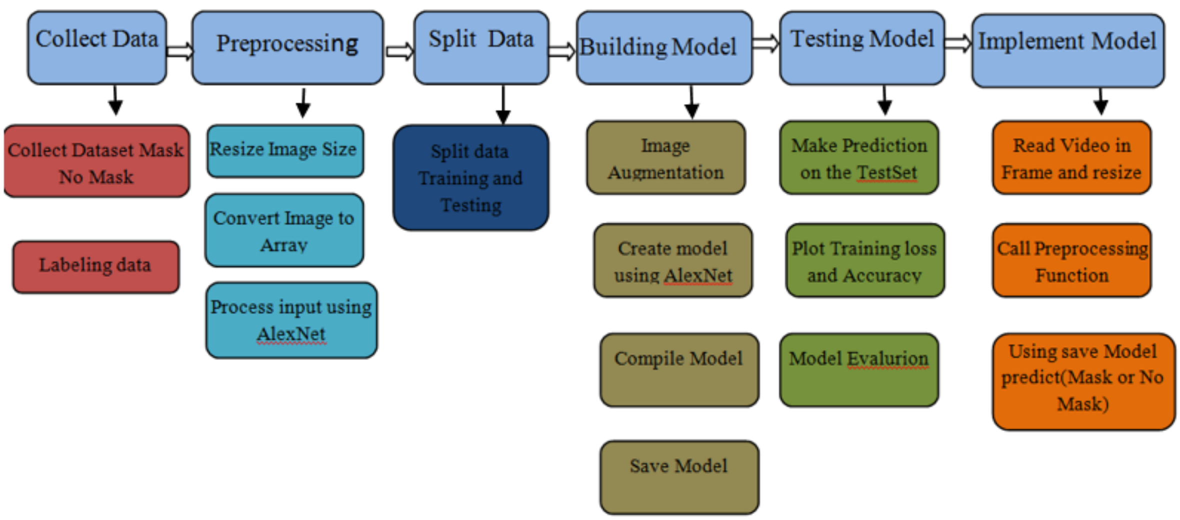

Fig. 3. Sequence of Steps for the Model

3.1 IoT Module

The main function of the IoT module is to receive the photos from the cameras installed in the smart city then using IoT devices these photos will be forwarded to our proposed model which consists of the preprocessing unit and CNN Model Biswal [17]. The following IoT devices are used for exploring the effectiveness of the proposed model.

3.1.1 Raspberry Pi

Raspberry Pi is a very small in size computer which can be connected to a monitor, keyboard and mouse. It is very small in size and explores the computing ability with the language like python. Figure 4 shows Raspberry Pi.

Fig. 4. Raspberry Pi

3.1.2 Pi camera

Pi-Camera is used in Raspberry Pi and is capable of taking image and video unless light. It connects with Raspberry pi using MIPI CSI-3 protocol that is primarily used for the transmission of images and videos between hosts and cameras in a UniPro-based M-PHY device network that is multilayered and peer-to-peer. Image processing, machine learning, and surveillance projects typically make use of it. The payload of this type of camera is extremely low so it is also used in surveillance drones. Figure 5 shows a Pi camera that is supposed to be connected with the Raspberry Pi device.

Fig. 5. Pi Camera

3.1.3 Temperature sensor

The Figure 6 shows the temperature sensor (MLX90614) which is a non-contact and an infrared temperature reader that reads a human's body temperature in a contactless manner. It is capable of handling temperatures of up to -19 to 130°C, but it's recommended you keep it below 100°C (210°F) with a solution achieved up to 0.15°c. Making use of the noise-reducing amplifier on the thermometer, a 17-bit ADC, and strong DSP unit that is significantly efficient in producing more accuracy.

Fig. 6. Temperature sensor

3.2 Image Preprocessing

Before images are processed in CNN model, all the images should be processed. As we are using PyTorch so all images were transformed into tensors. Thereafter for faster calculation the tensor list was converted to NumPy array Mallick [14]. After the model had been trained, the cycle of data augmentation was used to improve accuracy.

3.2.1 Data augmentation

Data augmentation techniques are those techniques that can be used to enhance the data size by virtue of incorporating some changes copies of previously existing data or recently created manufactured data from existing data [2].

The image is transformed in a number of ways, including: geometric transformations like flipping, rotating, scaling, cropping, and translation, as well as colour casting, changing brightness, and introducing motion.

In this step we have to transform all images to a specific format like grey-scaling, resize, centre -crop, and normalize. We have taken all images into size 256 to maintain uniformity of the input images, centre-crop 224 and normalize all images. The augmented images are shown in Figure 7.

Fig. 7. Augmented data

3.3 CNN Model
3.3.1 Transfer learning approach using AlexNet

Transfer learning with some adoption can be repurposed for another connected task rather than the original task itself. It is typically utilized whenever sufficient training samples are not available in order for training the model, like classification of medical images for various diseases. It is specifically true for models of various deep neural networks (DNN), that do not need to be trained on a significantly large number of performance parameters. With the help of transfer learning, the model parameters get off to a perfect beginning with presumably acceptable initial values which should only to be tweaked a little to make them better fit the new task.

Pre-trained model can be implemented for a variety of tasks in two primary ways. The first method uses a classifier that has been trained on top of the pre-trained here model in order for performing the classification and at the same time, it can serve as a feature extractor as well. The second method involves fine-tuning the new task for the entire network or a subset of it. As a result, the weights of one of the parameters of the pre-trained model serve as the new primary values and are updated during the training phase [15].

Eight layers with learnable parameters were used in this AlexNet transform learning experiment. The model is made up of 5 convolution layers along with3 fully-connected layers. By virtue of using the ReLU activation function, this model speeds up the process of training nearly six times and uses dropout layers (0.5) to reduce Overfitting. In Table 1, the detailed structure of the model along with various parameters of Alexnet has been depicted.

Table 1
Alexnet Model Characteristics

Layer	Filters/Neurons	Filter Size	Stride	Padding	Size of feature map	Activation
Input	None	None	None	None	227*227*3	None
Conv1	96	11*11	4	None	55*55*96	ReLU
MaxPool 1	None	3*3	2	None	27*27*96	None
Conv 2	256	5*5	1	2	27*27*256	ReLU
Max Pool 2	None	3*3*2	None	None	13*13*256	None
Conv 3	384	3*3	1	1	13*13*384	ReLU
Conv 4	384	3*3	1	1	13*13*384	ReLU
Conv 5	256	3*3	1	1	13*13*256	ReLU
Max Pool 3		3*3	2	None	6*6*256	None
Dropout 1	Rate=0.5	None	None	None	6*6*256	None
Fully connected1	None	None	None	None	9216	ReLU
Dropout 2	Rate=0.5	None	None	None	4096	None
Fully connected2	None	None	None	None	4096	ReLU
Fully connected3	None	None	None	None	1000	Softmax

Figure 8 represents the pictorial architecture of Alexnet with details of each layer.

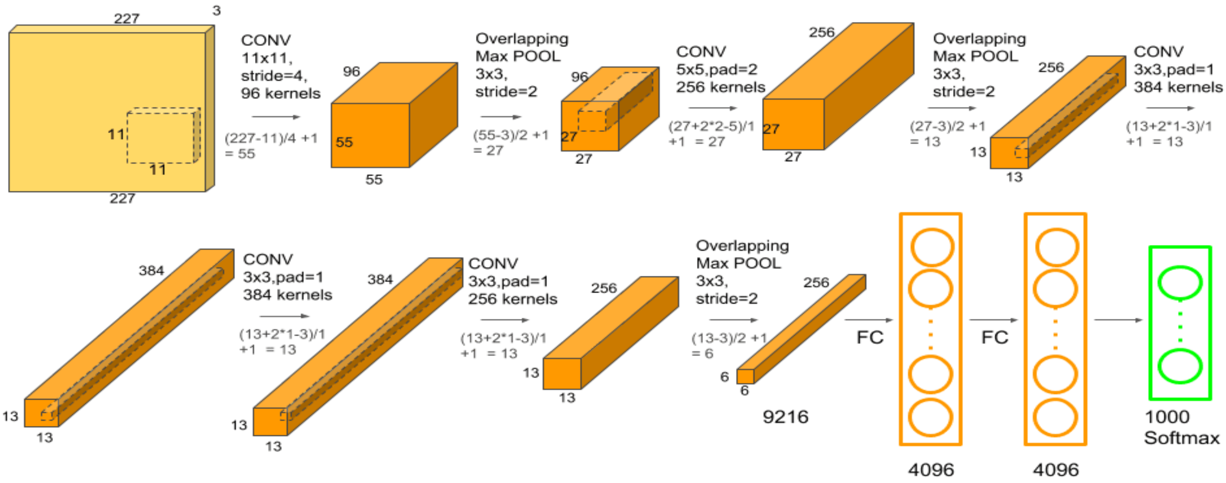

Fig. 8. Alexnet Architecture

3.3.2 Model architecture

The best learning model for image classification is CNN (Convolutional Neural Network). Various convolution layers make up the hidden layers, which learn reasonable filters for the extraction of significant features from the samples. Different dense neural networks use CNN's extracted features for classification purposes. The Overfitting of the network is to be managed by the dropout layer by virtue of dropping out the units. Consequently, the identification of various classes is conducted by a dense layer that comprises of two neurons.

3.3.3 Convolutional Layer

The central component of the CNN is the Convolution Layer. Convolution refers to the process by which two mathematical functions are combined to produce the third function. A sliding window strategy's convolution function, which aids in image feature extraction P. Nagrath *et al.*, [3]. The creation of feature maps is made easier by this convolution function. We will obtain the output C as shown in Eq. (1) when the Convolutional kernel, B, and one input image matrix, A, are combined:

$$C(T) = (A * B)(x) = \int_{-\infty}^{+\infty} A(T) \times B(T - X)dt \quad (1)$$

3.3.4 Pooling layer

By allowing for a smaller input matrix having sacrificed most of the features, the pooling activities speed up computations. The following are some examples that refer to pooling operations that are applicable to polling operations:

i. Max Pooling: Max pooling is that pooling operation which takes the maximum value that is currently available within a specific chosen region in which the kernel is present, and that uses that value to determine the value of the output matrix for that cell.
ii. Average Pooling: Average pooling is that pooling operation which takes up the average value of a specific chosen region in which the kernel is currently present and it represents the value of the output matrix for that particular cell.

iii. Min Pooling: Min pooling is the operation of pooling that takes the minimum available value from a specific chosen region that currently holds the kernel and that uses that value in order for determining the value of the output matrix or that particular cell.

3.3.5 Dropout layer

By eliminating arbitrary biased neurons from the model, this reduces the possibility of over-fitting during training. These neurons can be found in both visible and hidden layers P. Nagrath *et al.*, [3]. By changing the dropout proportion, a neuron's probability of being dropped can be altered.

3.3.6 Non-linear layer / activation function

The consequent value of the output from the hidden layer is calculated with the help of the activation functions making use of the function "sum of the product" taking into consideration the weights as well as inputs which will be added with bias. Then, this calculated value would represent the input to the subsequent layer Behera [16]. The proposed model makes use of two activation functions that are detailed below.

i. Softmax
Softmax is the type of activation function the value of which lies between 0 and 1 and that is capable of standardizing the input values in such a way so as to make the sum of the output values equal to 1. This activation function decides the probability of getting true or false for each of the classes.
If Input x is $x = \begin{bmatrix} 1.2 \\ 0.7 \\ 0.6 \end{bmatrix}$, then $F(x) = \begin{bmatrix} 0.48 \\ 0.28 \\ 0.24 \end{bmatrix}$, or in other words, if x represents the vector of inputs, then the value of Softmax function F(x) can be calculated using Eq. (2):

$$F(X_j) = \frac{x_j}{\sum_{i=1}^{n} x_i} \tag{2}$$

where n is the number of classes and x_j is the j^{th} value of vector.

ii. Rectified linear unit (ReLU)
ReLU represents an activation function that is piecewise linear, would directly produce the input if it its value is greater than zero and would produce zero otherwise as depicted below.

$$F(x) = \begin{cases} x & \text{if } x>0; \\ 0 & \text{if } x<=0; \end{cases}$$

The pre-trained AlexNet model is fine-tuned with the help of PyTorch. The convolution 2D layers of AlexNet typically have 256 nodes. The results of this study are two layers that have been fine-tuned for connections with 9216 nodes to 256 nodes and then 256 nodes to two nodes. Thus, the resulting here two classes are with face mask and without face mask. Occurrence of Overfitting is facilitated by a model learning the undesirable patterns occurring during the training tests. Consequently, the accuracy of testing decreases whereas the training accuracy rises. To prevent Overfitting, the model's training includes a 0.5 dropout., and "Adam" is used as the optimizer and "CrossEntropyLoss" as the loss parameters.20 epochs are used for the training. Sharifah Noha Zahirah Syed Abdul Nasir *et al.,* [22] proposed a method for finding the non-linear calculation using Graphical User Interface (GUI). They used MATLAB R2020b software for this work.

iii. Dataset Collection
Dataset collected thereby from the repository of github [2] has been used for the purposes of training as well as testing. There are a total of 1918 images of people without a mask and 1914 images of people wearing masks in the dataset. For testing purposes 15%, training purposes 80%, and for validation purposes 5% images of each class are used. Figure 9 shows the dataset's splitting, which is plotted in Python.

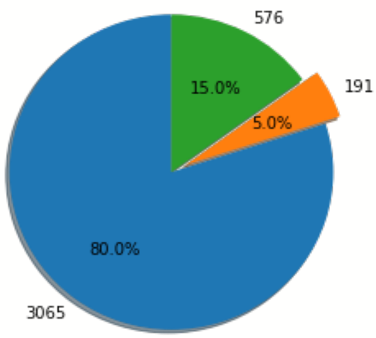

Fig. 9. Dataset description

4. Results and Discussion

During the CNN model's training, the 99.306 percent accuracy on tests was achieved. Figure 10 displays the confusion matrix from where we can calculate accuracy, precision and F-score which is given at Table 1 and Table 2. The sizes of the batches having been set to 32 and 20 iterations for the epochs, we have implemented VGG16, MobileNet and AlexNet.

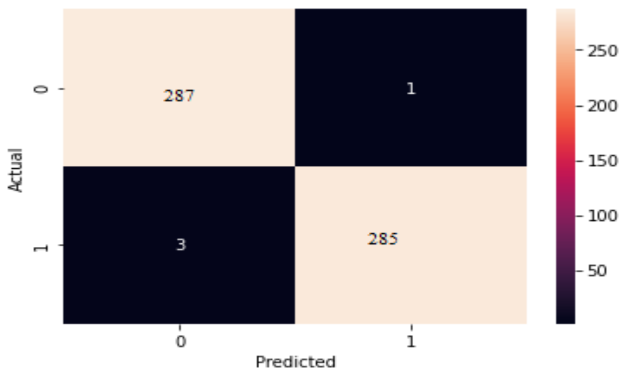

Fig. 10. The confusion matrix of the proposed model

AlexNet tends to outperform the other two specified models so far as the accuracy as well as loss are concerned, as shown in Figure 11 and Figure 12 shows performance of the testing results by virtue of visualization taking into consideration the level of accuracy as well as loss.

Fig. 11. The Loss and Epoch graph

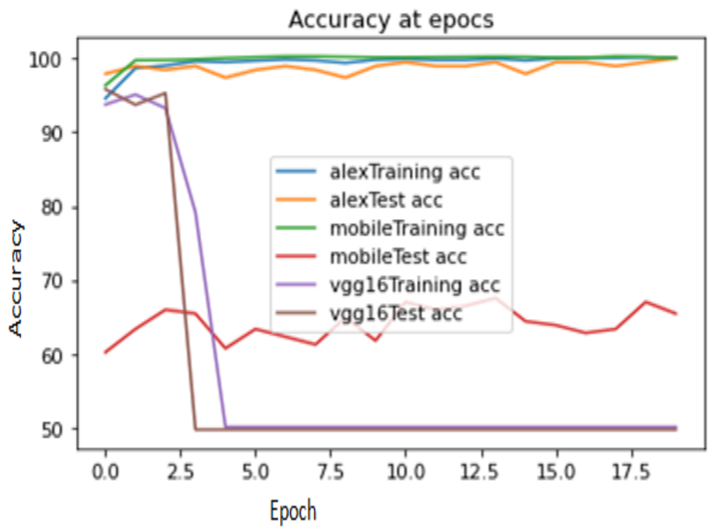

Fig. 12. The Accuracy curve of training and validation dataset used in proposed model for All model

Again, the graph plot of the accuracy and loss for AlexNet is illustrated in Figure 13. The following is a description of the accuracy of the analysed performance

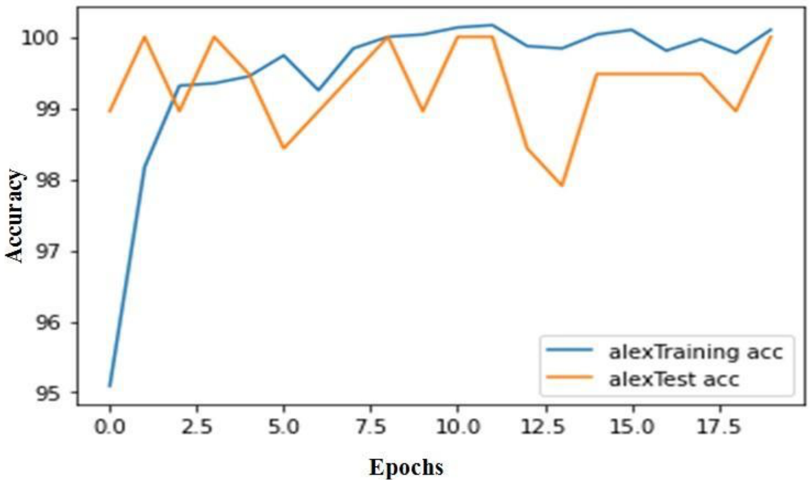

Fig. 13. The accuracy curve of training and validation dataset used in proposed model (AlexNet)

Figure 14 presents “Receiver Operating Characteristic (ROC)” curve of this proposed model. The expectation capacity of the classifier at various limits will be shown by the ROC curve. Taking two parameters we can plot the ROC curve; and the parameters are the true positive rate (TPR) and the other refers to the false positive rate (FPR) that have been measured by virtue of using Eq. (5) and Eq. (6) respectively. FPR and different thresholds are used to measure TPR, and the obtained thus values are plotted in the form of ROC curves. The performance of the binary classifier for each and every possible threshold is measured by the area under the curve (AUC) in the context of ROC Rath, M *et al.,* [19]. As it can be noticed, the value of the AUC value lies between 0 and 1. A model's AUC is one when it is 100% accurate, and it is zero when it is 100% incorrect. Our classifier model achieved an AUC of 0.993, indicating that it is a competent classifier.

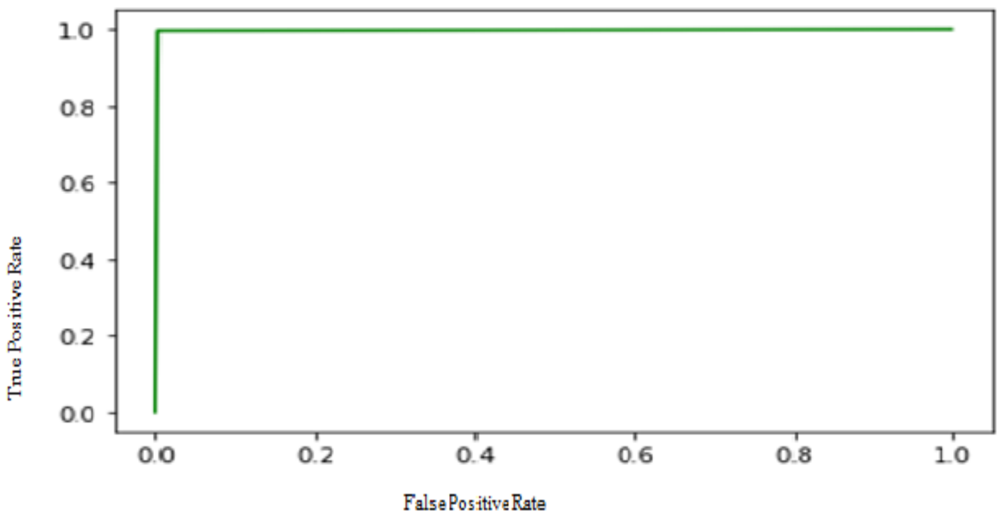

Fig. 14. The Receiver Operating Characteristic(ROC) curve

$$precision = \frac{N(TP)}{N(TP)+N(FP)} \quad (3)$$

$$\mathrm{Re}\,cal = \frac{N(TP)}{N(TP)+N(FN)} \quad (4)$$

$$F_1 Score = 2\,\frac{precission*recall}{precission+recall} \quad (5)$$

$$Accuracy = \frac{N(TP) + N(TN)}{N(TP) + N(TN) + N(FP) + N(FN)} \tag{6}$$

where total number of true positives that is represented by N(TP) and the total number of false positives is denoted by N(FP), the total number of true negatives has been represented by N(TN) and the total number of false negatives has been denoted by N(FN). By taking the average of all these measures across the two classes, a total measure of the algorithm can be calculated and computing it for each class. The analysis reveals that the developed MASK classification framework had a classification accuracy of 99.306 percent on average. 285 out of the 288 images that represent With Mask have their classifications correct, while the remaining three images have their classifications incorrect. The remaining 1 image is incorrectly classified as With Mask, while 287 of the 288 images Without Mask are correctly classified. The confusion matrix that was generated when the developed here model was validated by virtue of using the developed thereby dataset has been depicted in Table 2.

Table 2
Confusion Matrix of the Proposed Algorithm

True Class	Predicted	
	Abnormal	Normal
Abnormal (288)	287	1
Normal (288)	3	285

The parameters such as precision, recall and F-score are shown in Table 3. The With Mask category has the highest precision, at 100%, followed by Without Mask, which has a precision of 100%. The proposed here algorithm demonstrates 99.3 percent of precision on average. With 100%, the abnormal class achieves the highest recall; Normal 99 percent recall is the lowest, while 99 percent recall is the average. Lastly, the value of the F-score has been found to be 99% on an average, with Normal at its highest 99% and Abnormal at its lowest 99%. Some energy is needed by cameras to send photos and emails to appropriate authorities using this proposed model Panda, N. *et al.,* [18].

Table 3
Performance Analysis of the Proposed Algorithm for Detection and Classification of Face Mask Images

Class	Precision	Recall	$F1_{score}$	Support
Abnormal	1	0.99	0.99	288
Normal	0.99	1	0.99	288
Average	0.99	0.99	0.99	576

$$\text{True Positive Rate(TPR)} = \frac{TruePositive}{TruePositive + FalseNegative} \tag{7}$$

$$\text{False Positive Rate(FPR)} = \frac{FalsePositive}{TrueNegative + FalsePositive} \tag{8}$$

Before the face detection algorithm is activated, the model examines the video frame by frame. The caffe model is used to identify faces in this case Routray, P. K *et al.,* [21]. The image will be processed following face detection, including resizing and converting to an array for execution Reddy, D.S. *et al.,* [20]. After that, the processed image is given to the proposed model as an input, and the video frames are also labelled. In addition to the predicted percentage, the model's output predicts whether a face mask is on a person.

Figures 15, 16, 17 and 18 show a variety of test results regarding the model's capability to identify people who are and those who are not wearing facemasks. The image in Figure 15 and Figure 17 shows the result of identification of an individual without wearing a face mask, as depicted in the images, with an accuracy of 100%. In Figure 17 as it can be noticed, an individual appears to have put on a face mask that neither covers the nose nor the face completely. Figure 18 shows a person who has not put on a face mask with 78.40% of accuracy. Similarly, Figure 16 depicts a facemask-wearing individual with a 100% accuracy.

Fig. 15. Person with No-Mask 100%

Fig. 16. Person with Mask 100%

Fig. 17. Person with No-Mask 100%

Fig. 18. Person with No-Mask 78.40%

The proposed model implemented will generate an automatic email alert together with a photo of the person to the competent authority for initiating action if the percentage of wearing a mask is less than 80%. We use twilio [21] for sending email in the proposed algorithm which is displayed in Figure 19.

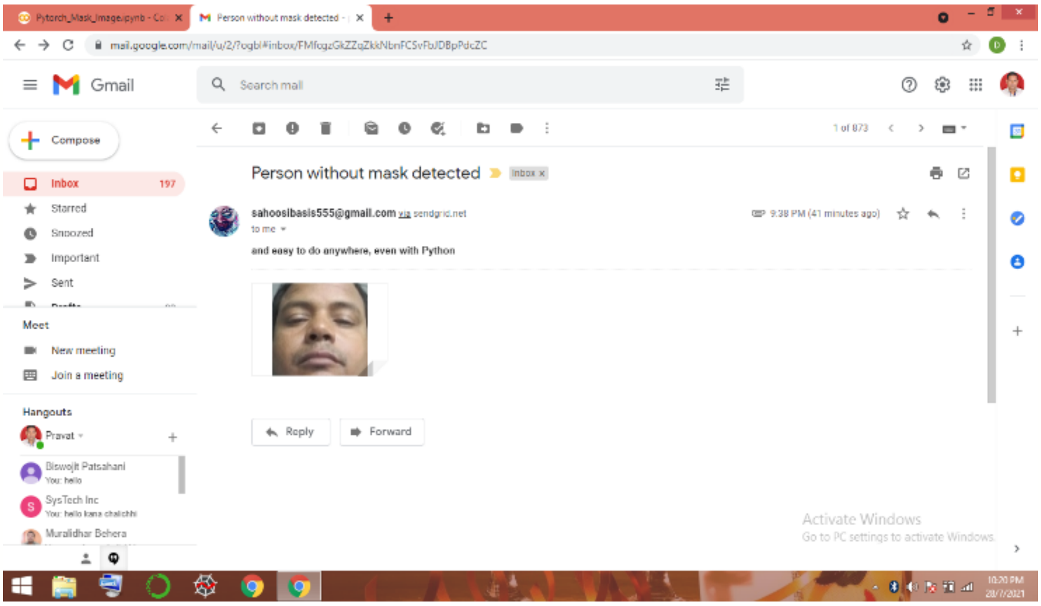

Fig. 19. Email to the authority along with photo

Table 4
Training and Validity Accuracy and Loss

Epoch	Loss	Accuracy	Val_Loss	Val_Accuracy
0	0.37213081	95.296	0.34442878	99.375
1	0.31393072	98.224	0.34453875	99.375
2	0.31359872	99.079	0.31326306	96.875
3	0.31332362	98.882	0.31330186	100.000
4	0.31488904	99.671	0.31455082	100.000
5	0.31326181	99.243	0.31339660	99.375
6	0.31327391	99.375	0.31326175	99.375
7	0.31326264	99.704	0.31326190	100.000
8	0.31446686	99.868	0.31326169	99.375
9	0.31331947	99.211	0.31328073	99.375
10	0.31333685	99.243	0.31326169	100.000
11	0.32577586	99.671	0.31326371	99.375
12	0.34451172	99.737	0.31332663	99.375
13	0.31327507	99.868	0.31326169	99.375
14	0.31326234	99.375	0.31326586	100.000
15	0.31326199	99.638	0.31326234	99.375
16	0.31333774	99.704	0.31326175	100.000
17	0.31326240	99.737	0.31326169	99.375
18	0.31326303	99.572	0.31326169	99.375
19	0.31326172	99.474	0.31456795	99.375

Table 5
External Validation

Author	Model	Accuracy (%)
Bosheng Qin and Dongxiao Li	SRCNet	98.70
S. A. Sanjaya and S. A. Rakhmawan,	MobileNetV2	92
Gabriel T S Draughon *et al.,*	R-CNN	96
Mohamed Loey *et al.,*	ResNet 50	81
S. V. Militante and N. V. Dionisio,	VGG-16	96
G K Jakir Hussain *et al.,*	Proposed Method CNN	91.11
Preeti Nagratha *et al.,*	MobileNetV2	92.64
OurModel	**Proposed CNN Model**	**99**

5. Conclusion

The proposed approach has been conducted on a two-fold basis, it is involved in a single IoT framework, each part of the work has been found to demonstrate better results in comparison with the earlier models. Using deep learning methods for prevention, the detection procedure is approximately 100% accurate. Additionally, the internet communication is verified. For the mask similar to face is not considered here and kept for future work.

Acknowledgement

This research was not funded by any grant.

References

[1] WHO EMRO. “About COVID-19 | COVID-19 | Health topics.” http://www.emro.who.int/health-topics/coronavirus/about-covid-19.html

[2] Balajisrinivas. “Detecting face masks using Python, Keras, OpenCV on real video streams” https://github.com/balajisrinivas/Face-Mask-Detection

[3] Nagrath, Preeti, Rachna Jain, Agam Madan, Rohan Arora, Piyush Kataria, and Jude Hemanth. "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2." *Sustainable cities and society* 66 (2021): 102692. https://doi.org/10.1016/j.scs.2020.102692

[4] Sanjaya, Samuel Ady, and Suryo Adi Rakhmawan. "Face mask detection using MobileNetV2 in the era of COVID-19 pandemic." In *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, pp. 1-5. IEEE, 2020. https://doi.org/10.1109/ICDABI51230.2020.9325631

[5] Militante, Sammy V., and Nanette V. Dionisio. "Real-time facemask recognition with alarm system using deep learning." In *2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC)*, pp. 106-110. IEEE, 2020. https://doi.org/10.1109/ICSGRC49013.2020.9232610

[6] Hussain, GK Jakir, R. Priya, S. Rajarajeswari, P. Prasanth, and N. Niyazuddeen. "The face mask detection technology for Image analysis in the covid-19 surveillance system." In *Journal of Physics: Conference Series*, vol. 1916, no. 1, p. 012084. IOP Publishing, 2021. https://doi.org/10.1088/1742-6596/1916/1/012084

[7] Loey, Mohamed, Gunasekaran Manogaran, Mohamed Hamed N. Taha, and Nour Eldeen M. Khalifa. "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection." *Sustainable cities and society* 65 (2021): 102600. https://doi.org/10.1016/j.scs.2020.102600

[8] Bosheng, Qin, and Dongxiao Li. "Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network." *Mdpi Sens* (2020): 1-23.

[9] Das, Abhishek, and Mihir Narayan Mohanty. "Covid-19 detection from x-ray images using convolutional neural network." *International Journal of Advanced Science and Technology* 29, no. 8 (2020). https://doi.org/10.21203/rs.3.rs-51360/v1

[10] Mohanty, Mohan Debarchan, Pradeep Kumar Mallick, and Mihir Narayan Mohanty. "Design of an Intelligent Controller of a Self-Tuned Quadcopter for IoT-Based Application on COVID Infected Area Sanitization." *IEEE Internet of Things Magazine* 4, no. 3 (2021): 24-29. https://doi.org/10.1109/IOTM.1100.2000068

[11] Das, Abhishek, Saumendra Kumar Mohapatra, Asit Subudhi, and Mihir Narayan Mohanty. "Classification of COVID-19 infectedX-ray image using deep learning techniques." *Annals of the Romanian Society for Cell Biology* (2021): 2736-2743.

[12] Patra, Gyana Ranjan. "Prediction of COVID Patients using Gated Recurrent Unit Neural Networks." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12, no. 9 (2021): 2333-2340.

[13] de Albuquerque, Victor Hugo C. *Hybrid Artificial Intelligence and IoT in Healthcare*. Edited by Akash Kumar Bhoi, Pradeep Kumar Mallick, and Mihir Narayana Mohanty. Springer, 2021.

[14] Mallick, Pradeep Kumar, Bhabani Sankar Satapathy, M. N. Mohanty, and S. Saravana Kumar. "Intelligent technique for CT brain image segmentation." In *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*, pp. 1269-1277. IEEE, 2015. https://doi.org/10.1109/ECS.2015.7124789

[15] Behera, Shreetam, Miihir Narayan Mohanty, and Srikanta Patnaik. "A comparative analysis on edge detection of colloid cyst: a medical imaging approach." In *Soft computing techniques in vision science*, pp. 63-85. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. https://doi.org/10.1007/978-3-642-25507-6_7

[16] Jyoti, Amlan, Mihir Narayan Mohanty, and Mallick Pradeep Kumar. "Morphological based segmentation of brain image for tumor detection." In *2014 International Conference on Electronics and Communication Systems (ICECS)*, pp. 1-5. IEEE, 2014. https://doi.org/10.1109/ECS.2014.6892750
[17] Biswal, Anil Kumar, Debabrata Singh, Binod Kumar Pattanayak, Debabrata Samanta, and Ming-Hour Yang. "IoT-based smart alert system for drowsy driver detection." *Wireless communications and mobile computing* 2021 (2021): 1-13. https://doi.org/10.1155/2021/6627217
[18] Panda, Niranjan, Prabhat Ku Sahu, Manoranjan Parhi, and Binod Ku Pattanayak. "A survey on energy awareness mechanisms in aco-based routing protocols for manets." *Intelligent and Cloud Computing: Proceedings of ICICC 2019, Volume 1* (2020): 791-800. https://doi.org/10.1007/978-981-15-5971-6_85
[19] Rath, Mamata, and Binod Pattanayak. "Technological improvement in modern health care applications using Internet of Things (IoT) and proposal of novel health care approach." *International Journal of Human Rights in Healthcare* 12, no. 2 (2018): 148-162. https://doi.org/10.1108/IJHRH-01-2018-0007
[20] Reddy, D. Santhosh, and Pachamuthu Rajalakshmi. "A Novel web application framework for ubiquitous classification of fatty liver using ultrasound images." In *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, pp. 502-506. IEEE, 2019. https://doi.org/10.1109/WF-IoT.2019.8767283
[21] Routray, Pravat Kumar, Binod Kumar Pattanayak, and Mihir Narayan Mohanty. "An IoT Application Framework Using Deep Learning for Face Mask Detection." *Specialusis Ugdymas* 1, no. 43 (2022): 5450-5464.
[22] Ab Wahab, Nurul Ain, and Mohd Agos Salim Nasir. "Graphical User Interface for Solving Non-Linear Equations for Undergraduate Students." *International Journal of Advanced Research in Future Ready Learning and Education* 30, no. 1 (2023): 25-34.
[23] Ha, Chin Yee, Terh Jing Khoo, and Zheng Yik Koo. "Current status of green building development in Malaysia." *Progress in Energy and Environment* (2023): 1-9. https://doi.org/10.37934/progee.25.1.19